

Dritter Zwischenbericht zur Studie: Körpergröße als Indikator sozialer Ungleichheit

Mumdzhiev, Milko

Preprint / Preprint

Zwischenbericht / interim report

Empfohlene Zitierung / Suggested Citation:

Mumdzhiev, M. (2011). *Dritter Zwischenbericht zur Studie: Körpergröße als Indikator sozialer Ungleichheit*. (Nürnberger Beiträge zur Sozial- und Wirtschaftsforschung, 04/11). Nürnberg. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-250858>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see: <https://creativecommons.org/licenses/by-nc-nd/4.0>

Dritter Zwischenbericht zur Studie: Körpergröße als Indikator sozialer Ungleichheit

Milko Mumdzhev

Zusammenfassung:

Anhand des ALLBUS 2008 wird die Variable Körpergröße der befragten Person als Indikator sozialer Ungleichheit getestet. Korrelationsmessungen und lineare Regressionen, Logitmodelle (binär, ordinal, multinomial), Diskriminanzanalysen, und eine Lebensstilbetrachtung anhand Clusteranalysen plus Varianzanalysen sollen die Bedeutung des Merkmals Körpergröße der Person als abhängige und unabhängige Variable in diversen Zusammenhängen der Sozialstrukturanalyse aufzeigen.

Einleitung

In diesem dritten Zwischenbericht zur Studie Körpergröße als Indikator sozialer Ungleichheit werden nacheinander Korrelationen der Körpergröße mit stetigen Variablen des ALLBUS 2008 Datensatzes berichtet, lineare Regressionen mit Körpergröße als abhängige Variable, dazu binäre und multinomiale logistische Modelle mit Körpergröße als unabhängige Variable geliefert. Desweiteren werden mehrere Lebensstiloperationalisierungen vorgenommen und durch Varianzanalysen, multinomiale logistische Regressionen (und Diskriminanzanalysen) mit Körpergröße in Verbindung gesetzt.

Um SPSS-Befehle wie CLUSTER oder GENLIN, die die Verwendung von Gewichtung kaum oder gar nicht unterstützen, dennoch im gesamtdeutschen Kontext anwenden zu können, wurde aus dem ursprünglichen ungewichteten ALLBUS 2008 Datensatz (mit dem Sample-Befehl) eine Stichprobe gezogen, um die überrepräsentierte Bevölkerung der neuen Bundesländer zu reduzieren.

2392 und 1077 (bzw. 69,0% und 31,0%) Befragte in den alten und neuen Bundesländern werden durch Gewichtung mit v_{792} zu (gerundeten) 2830 und 639 (bzw. 81,6% und 18,4%) Fällen der entsprechenden Gruppen. Nachdem aus den ursprünglichen 1077 Befragten der neuen Bundesländern 540 Fälle gezogen werden (knapp die Hälfte), bildeten die verbleibenden 2932 Fälle den vorliegenden gesamtdeutschen Datensatz.

Für den gewichteten Originaldatensatz erhält man für das Merkmal Körpergröße:

	N	Min.	Max.	Mittel	SA	Schiefe	Kurtosis
Mann	814	158	206	177,69	6,959	,182	,042
Frau	854	146	184	165,14	6,196	-,011	-,278

Die selben deskriptiven Statistiken am reduzierten Datensatz lauten:

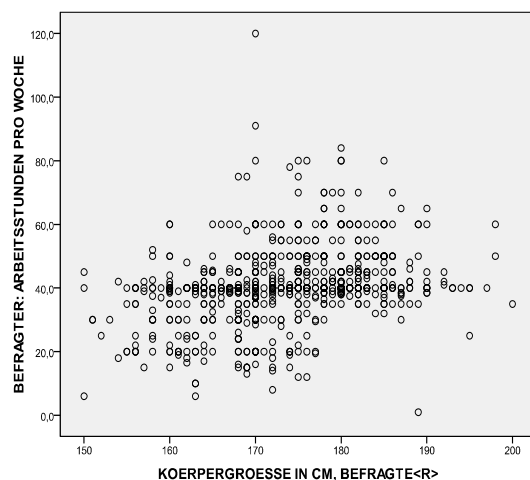
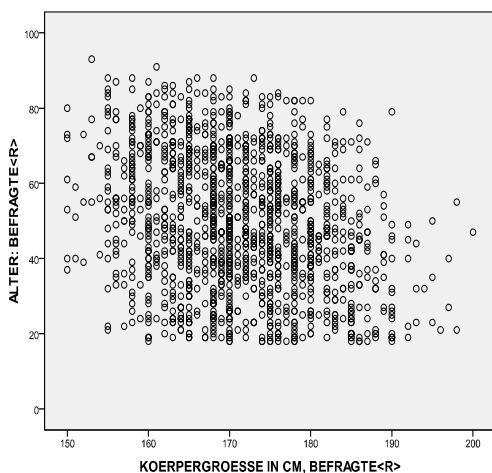
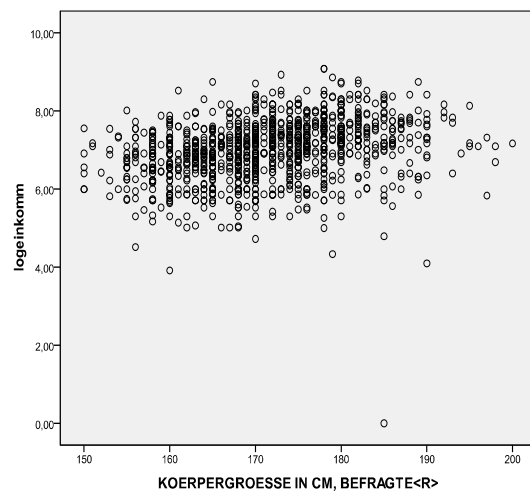
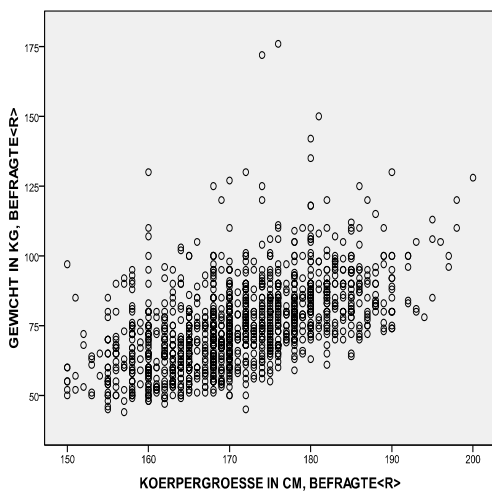
	N	Min.	Max.	Mittel	SA	Schiefe	Kurtosis
Mann	690	158	200	177,64	6,869	,206	-,047
Frau	728	150	184	165,11	6,167	,023	-,332

Korrelationen nach Pearson

Stetige Variablen des ALLBUS 2008 wurden mit Körpergröße des Befragten korreliert, wobei der Fall 751 ($v2=751$) jedesmal (auch in den weiteren Rechnungen) ausgeschlossen wurde (sein Z-Score betrug 3,1 bzw. 3,3 nur auf Männer bezogen; was aber praktisch keine Rolle spielte). Variable v388, das Nettoeinkommen erfassend, wurde sowohl in logarithmierter Form, als auch in der ursprünglichen angegeben.

Die Verteilungen der metrischen Variablen waren selten normal; Einkommen, Prestigewerte etc. konnten nach einer Logarithmierung (und weiteren Manipulationen) zu Analysezwecken „hingebogen“ werden; so ging auch in die folgende Regressionsrechnung das logarithmierte Einkommen ein.

Betrachtet man Streudiagramme zwischen Körpergröße und anderen (nicht transformierten) stetigen Variablen, wird kaum das ovale Zigarrenmuster als Zeichen von Linearität ausgemacht. Am ehesten kann man Gewicht oder das **logarithmierte** Einkommen zitieren, nicht aber das z.B. das Alter.



Nimmt man die SPSS-Prozedur CURVEFIT zwecks Prüfung diverser Zusammenhänge zwischen Körpergröße und anderen Variablen, werden lineare Zusammenhänge gegenüber anderen in Termen von R^2 hintangestellt (selbst bei Gewicht).

Die (absolut) größten pearsonschen Korrelationskoeffizientenwerte wurden bei Alter, Gewicht und den Prestigeangaben (auch der Eltern) gefunden. Auf dieser Ebene der Analyse waren bei Frauen die Korrelationen mit Prestigeangaben der Mütter größer als die der Väter, und umgekehrt.

Die erste Zeile der Tabellen enthält die Variablenlabels, die zweite den Wert des Korrelationskoeffizienten, die dritte das Signifikanzniveau, die vierte und letzte Zeile liefert die Fallzahl. Gelb markiert sind die signifikanten Ergebnisse. Spearmans Korrelationskoeffizienten, sofern auf dem 0,05 Niveau im **Gegensatz zu Pearsons r** nicht oder doch signifikant ausgegeben, sind durch **Fettmarkierung** angedeutet.

Frauen

ATTRAKTIVITÄT DES BEFR., INTERVIEW- START	FERNSEHGE- SAMTDAUER PRO TAG IN MINUTEN	LOGARITH- MIERTES EINKOMMEN (V388)	ALTER: BEFRAGTE<R>	BEFR.: JETZIGES TREIMAN- PRESTIGE I68
,148**	-,113**	,139**	-,281**	,108
,000	,003	,001	,000	,064
728	710	547	725	292

Für das Merkmal Attraktivität der Befragten (in der ersten Spalte, ordinal skaliert mit 11 Ausprägungen) wird Spearmans Korrelationskoeffizient extra erwähnt: 0,162 (hochsignifikant auf 0,000).

BEFR.: JETZIGES MAGNITUDE- PRESTIGE I68	BEFR.: ISEI NACH GANZEBOOM I68	BEFR.: JETZIGER BERUF; SIOPS I88	BEFR.: JETZIGES MAGNITUDE- PRESTIGE I88	BEFR.: ISEI NACH GANZEBOOM I88
,077	,076	,063	,091	,097
,192	,232	,293	,126	,105
292	246	283	283	283

BEFRAGTER: ARBEITS- STUNDEN PRO WOCHE	ANZAHL, BE- SCHÄFTIGTE B.D. ARBEITSSTELLE	BEFR.: NEBEN- ERWERB; ARBEITSSTD. PRO WOCHE	BEFR.: MAGNITUDE- PRES., LETZTER BERUF I68	BEFR.: ISEI GANZEBOOM, LETZTER BERUF I68
,149*	,124*	,071	,081	,099
,010	,043	,630	,132	,092
295	265	49	344	289

BEFR.: LETZTER BERUF; SIOPS I88	BEFR.: MAGNITUDE- PRES., LETZTER BERUF I88	BEFR.: ISEI GANZEBOOM, LETZTER BERUF I88	ANZ. BESCHÄFTIGTE, LETZTE ARBEITSSTELLE	DAUER <EHMALIGER> ARBEITSLOSIG- KEIT
,083	,099	,091	,163	-,085
,131	,070	,096	,532	,477
334	334	334	17	72

GEGENW. EHEPARTNER: ALTER	GEGENW.EHEP.: JETZIG.TREIMAN- PRESTIGE I68	GEGENW.EHEP.: JETZIG. MAGNITUDE- PREST.I68	GEGENW.EHEP.: ISEI NACH GANZEBOOM I68	GEGENW.EHEP.: SIOPS I88
-,257**	,129	,125	,168*	,139*
,000	,053	,060	,023	,040
395	227	227	184	219

GEGENW.EHEP.: ISEI NACH GANZEBOOM I88	GEGENW.EHEP.: JETZIG. MAGNITUDE- PREST.I88	LEBENS- PARTNER: ALTER	LEBENS- PARTNER: ISEI NACH GANZEBOOM I88	LEBENS- PARTNER: MAGNITUDE- PRESTIGE I68
,128	,126	-,131	,127	,009
,059	,061	,197	,306	,942
219	221	99	67	67

LEBENS- PARTNER: SIOPS I88	LEBENS- PARTNER: ISEI NACH GANZEBOOM I88	LEBENS- PARTNER: MAGNITUDE- PRESTIGE I68	VATER: TREIMAN- PRESTIGE I68	VATER: MAGNITUDE- PRESTIGE I68
,117	-,008	,010	,093*	,117**
,382	,952	,943	,024	,004
58	61	58	587	587

VATER: ISEI NACH GANZEBOOM I68	VATER: SIOPS I88	VATER: MAGNITUDE- PRESTIGE I88	VATER: ISEI NACH GANZEBOOM I88	MUTTER: TREIMAN- PRESTIGE I68
,122**	,096*	,132**	,123**	,143**
,008	,025	,002	,004	,008
472	551	557	551	339

MUTTER: MAGNITUDE- PRESTIGE I68	MUTTER: ISEI NACH GANZEBOOM I68	MUTTER: SIOPS I88	MUTTER: MAGNITUDE- PRESTIGE I88	BEFR.: NETTOEIN- KOMMEN, OFFENE ABFRAGE
,157**	,189**	,170**	,191**	,150**
,004	,002	,002	,001	,001
339	270	324	324	478

BFR.:NETTO-EINKOMMEN- <OFFENE+ LISTENANGABE>	HAUSHALTS- NETTOEIN- KOMMEN: OFFENE ABFRAGE	HAUSHALTS- EINKOMMEN <OFFENE+ LISTENANGABE>	ALTER, 1.KIND, AUSSER HAUS	ALTER, 2.KIND, AUSSER HAUS
,178**	,146**	,188**	-,206**	-,075
,000	,001	,000	,000	,295
547	498	603	328	197

ALTER, 3.KIND, AUSSER HAUS	GEWICHT IN KG, BEFRAGTE<R>	BODY-MASS- INDEX	TREIMAN- PRESTIGE I68, EINORD. NACH TERWEY	MAGNITUDE- PRESTIGE I68, EINORD. N. TERWEY
,036	,220**	-,180**	,157**	,144**
,735	,000	,000	,000	,000
91	681	681	682	682

ISEI GANZEBOOM I68, EINORDNUNG N. TERWEY	SIOPS I88, EINORDNUNG NACH TERWEY	MAGNITUDEPRESTIGE I88, EINORD. N. TERWEY	ISEI GANZEBOOM I88, EINORDNUNG N. TERWEY
,140**	,130**	,151**	,143**
,001	,001	,000	,000
573	661	662	661

Männer

Spearman's Koeffizient für Attraktivität lautet 0,212 (hochsignifikant auf 0,000).

Fettgedruckt sind auch hier diejenigen pearsonsche Koeffizienten, die sich hinsichtlich ihrer 0,05-Signifikanz bzw. Nichtsignifikanz von den Spearmanwerten unterscheiden (z.B. 0,08 und 0,02).

ATTRAKTIVITÄT DES BEFR., INTERVIEWSTAR T	FERNSEH- GESAMTDAUER PRO TAG IN MINUTEN	LOGARITH- MIERTES EINKOMMEN (V388)	ALTER: BEFRAGTE<R>	BEFR.: JETZIGES TREIMAN- PRESTIGE I68
,203**	-,095*	,081*	-,365**	,075
,000	,014	,049	,000	,138
689	667	585	688	390

BEFR.: JETZIGES MAGNITUDE- PRESTIGE I68	BEFR.: ISEI NACH GANZEBOOM I68	BEFR.: JETZIGER BERUF; SIOPS I88	BEFR.: JETZIGES MAGNITUDE- PRESTIGE I88	BEFR.: ISEI NACH GANZEBOOM I88
,078	,125*	,092	,139**	,162**
,125	,024	,075	,007	,002
390	326	378	382	378

BEFRAGTER: ARBEITS- STUNDEN PRO WOCHE	ANZAHL, BESCHÄFTIGTE B.D. ARBEITSSTELLE	BEFR.:NEBEN- ERWERB; ARBEITSSTD. PRO WOCHE	BEFR.: MAGNITUDE- PRES., LETZTER BERUF I68	BEFR.: ISEI GANZEBOOM, LETZTER BERUF I68
,031	,103*	,035	,051	,140*
,539	,047	,885	,427	,047
393	370	19	246	202

BEFR.: LETZTER BERUF; SIOPS I88	BEFR.: MAGNITUDE- PRES., LETZTER BERUF I88	BEFR.: ISEI GANZEBOOM, LETZTER BERUF I88	ANZ. BESCHÄFTIGTE, LETZTE ARBEITSSTELLE	DAUER <EHMALIGER> ARBEITSLOSIG- KEIT
,098	,111	,125	-,226	-,169
,132	,087	,055	,366	,170
236	240	236	18	67

GEGENW. EHEPARTNER: ALTER	GEGENW.EHEP.: JETZIG.TREIMAN PRESTIGE I68	GEGENW.EHEP.: JETZIG. MAGNITUDE- PREST.I68	GEGENW.EHEP.: ISEI NACH GANZEBOOM I68	GEGENW.EHEP.: SIOPS I88
-,254**	,040	,087	,039	,099
,000	,618	,280	,654	,217
413	157	157	136	157

GEGENW.EHEP.: ISEI NACH GANZEBOOM I88	GEGENW.EHEP.: JETZIG. MAGNITUDE- PREST.I88	LEBENS- PARTNER: ALTER	LEBENS- PARTNER: ISEI NACH GANZEBOOM I88	LEBENS- PARTNER: MAGNITUDE- PRESTIGE I68
,153	,156	-,359**	,287*	,230
,056	,052	,001	,037	,097
157	157	87	53	53

LEBENS- PARTNER: SIOPS I88	LEBENS- PARTNER: ISEI NACH GANZEBOOM I88	LEBENS- PARTNER: MAGNITUDE- PRESTIGE I68	VATER: TREIMAN- PRESTIGE I68	VATER: MAGNITUDE- PRESTIGE I68
,229	,232	,252	,116**	,131**
,102	,098	,072	,005	,001
52	52	52	585	585

VATER: ISEI NACH GANZEBOOM I68	VATER: SIOPS I88	VATER: MAGNITUDE- PRESTIGE I88	VATER: ISEI NACH GANZEBOOM I88	MUTTER: TREIMAN- PRESTIGE I68
,176**	,138**	,168**	,167**	,057
,000	,001	,000	,000	,321
469	544	555	544	301

MUTTER: MAGNITUDE- PRESTIGE I68	MUTTER: ISEI NACH GANZEBOOM I68	MUTTER: SIOPS I88	MUTTER: MAGNITUDE- PRESTIGE I88	BEFR.: NETTOEIN- KOMMEN, OFFENE ABFRAGE
,101	,101	,100	,179**	,122**
,080	,113	,089	,002	,007
301	250	291	291	496

BFR.:NETTO-EINKOMMEN <OFFENE+LISTENANGABE>	HAUSHALTS-NETTOEINKOMMEN: OFFENE ABFRAGE	HAUSHALTS-EINKOMMEN <OFFENE+LISTENANGABE>	ALTER, 1.KIND, AUSSER HAUS	ALTER, 2.KIND, AUSSER HAUS
,135**	,178**	,202**	-,214**	-,284**
,001	,000	,000	,000	,000
585	467	566	277	176

ALTER, 3.KIND, AUSSER HAUS	GEWICHT IN KG, BEFRAGTE<R>	BODY-MASS-INDEX	TREIMAN- PRESTIGE I68, EINORD. NACH TERWEY	MAGNITUDE- PRESTIGE I68, EINORD. N. TERWEY
-,283*	,359**	-,107**	,109**	,096*
,033	,000	,006	,005	,013
57	665	665	672	672

ISEI GANZEBOOM I68, EINORDNUNG N. TERWEY	SIOPS I88, EINORDNUNG NACH TERWEY	MAGNITUDEPRESTIGE I88, EINORD. N. TERWEY	ISEI GANZEBOOM I88, EINORDNUNG N. TERWEY
,154**	,120**	,155**	,173**
,000	,002	,000	,000
554	649	657	649

Pearsons Korrelationen für Unterpopulationen

Männer und Frauen wurden daneben unterschieden nach: alte und neue Bundesländer (AltB, NeuB), verheiratet (zusammenlebend, Verh), ledig, Angestellte (Anges), Arbeiter (Arbeit), ganztags erwerbstätig (GaEr), erwerbslos (Arblos), sowie mittlere Reife (MittR) und Hochschulreife (HochR). Tabelliert wurden Pearsons Korrelationen (samt Signifikanz und Fallzahl) zwischen Körpergröße und den Variablen:

Attraktivität zum Interviewstart (1.Spalte), Fernsehgesamtdauer pro Tag (2.Spalte), logarithmiertes Nettoeinkommen (3.Spalte), Alter (4.Spalte), des Vaters (5.Spalte) und der Mutter (6.Spalte) Magnitudeprestige (I88), Gewicht (7.Spalte) und Magnitudeprestige der Befragten (Einordnung nach Terwey, 8. Spalte) bzw. die Variablen v5, v12, ln (v388), v154, v351, v370, v631 und v785 des ALLBUS 2008. Spearmans Koeffizienten, sofern in ihrer Signifikanz unterschieden, sind durch Fettmarkierung angedeutet.

		ATTR	FERN	LOGEI	ALTER	VA I88	MU I88	GEWI	MA I88
Männer AltB	r	,188**	-,115**	,072	-,367**	,195**	,159*	,351**	,164**
	Sig.	,000	,007	,118	,000	,000	,028	,000	,000
	N	560	540	470	559	457	193	539	535
Frauen AltB	r	,150**	-,132**	,144**	-,305**	,124**	,187**	,221**	,172**
	Sig.	,000	,002	,003	,000	,008	,005	,000	,000
	N	583	567	423	580	450	227	548	523
Männer NeuB	r	,275**	-,024	,101	-,353**	,029	,227*	,397**	,093
	Sig.	,002	,790	,281	,000	,775	,025	,000	,307
	N	129	127	115	129	98	98	126	122
Frauen NeuB	r	,143	,020	,092	-,158	,140	,159	,236**	,050
	Sig.	,086	,815	,307	,058	,150	,119	,006	,556
	N	145	143	124	145	107	97	133	139
Männer Verh	r	,170**	-,195**	,304**	-,283**	,134*	,104	,425**	,190**
	Sig.	,000	,000	,000	,000	,013	,199	,000	,000
	N	421	415	367	421	340	153	405	403
Frauen Verh	r	,149**	-,133**	,194**	-,281**	,143*	,259**	,195**	,148**
	Sig.	,003	,009	,001	,000	,011	,001	,000	,004
	N	400	393	282	399	318	170	382	367
Männer Ledig	r	,208**	,009	-,102	-,336**	,106	,159	,273**	,128
	Sig.	,004	,900	,224	,000	,187	,110	,000	,088
	N	189	177	145	188	157	102	183	178
Frauen Ledig	r	,202*	-,103	,033	-,216**	,174	,162	,323**	,230**
	Sig.	,014	,220	,736	,009	,062	,140	,000	,007
	N	148	143	110	147	115	84	135	137
Männer Anges	r	,250**	-,051	,106	-,296**	,154*	,186	,445**	,028
	Sig.	,000	,454	,141	,000	,036	,071	,000	,685
	N	226	219	193	226	187	95	218	214
Frauen Anges	r	,165**	-,102*	,090	-,290**	,132*	,279**	,216**	,189**
	Sig.	,001	,047	,113	,000	,019	,000	,000	,000
	N	387	380	311	387	313	186	367	370
Männer Arbeit	r	,201**	-,142*	,023	-,387**	,034	,024	,359**	,083
	Sig.	,001	,022	,723	,000	,627	,799	,000	,181
	N	267	260	233	267	212	114	256	264
Frauen Arbeit	r	,077	-,073	,037	-,241**	,050	-,020	,238**	,017
	Sig.	,325	,354	,692	,002	,591	,867	,003	,832
	N	166	162	115	165	116	74	156	163

		ATTR	FERN	LOGEI	ALTER	VA I88	MU I88	GEWI	MA I88
Männer GaEr	r	,075	-,008	,014	-,203**	,090	,061	,366**	,140**
	Sig.	,146	,879	,798	,000	,106	,411	,000	,007
	N	382	368	326	382	323	185	371	369
Frauen GaEr	r	,069	-,071	,114	-,097	,092	,260**	,317**	,112
	Sig.	,347	,335	,151	,183	,261	,009	,000	,135
	N	190	186	161	189	150	99	171	179
Männer Arblos	r	,276**	-,137*	-,049	-,446**	,314**	,306**	,340**	,227**
	Sig.	,000	,024	,460	,000	,000	,003	,000	,000
	N	276	270	229	275	206	95	263	259
Frauen Arblos	r	,163**	-,064	,061	-,306**	,196**	,242**	,191**	,152**
	Sig.	,001	,217	,330	,000	,001	,004	,000	,005
	N	383	372	255	381	272	143	365	333
Männer MittR	r	,143*	-,068	-,005	-,312**	,040	,068	,418**	,012
	Sig.	,049	,359	,950	,000	,617	,508	,000	,867
	N	190	185	159	190	157	98	183	182
Frauen MittR	r	,178*	-,014	,100	-,135	,070	,074	,283**	,122
	Sig.	,011	,843	,213	,054	,371	,466	,000	,090
	N	204	202	156	204	168	100	188	193
Männer HochR	r	,055	-,095	,074	-,289**	,066	,226	,409**	-,040
	Sig.	,476	,239	,381	,000	,450	,066	,000	,615
	N	167	156	144	166	132	67	159	157
Frauen HochR	r	,118	-,133	,084	-,233**	,173	,290*	,336**	,152
	Sig.	,158	,122	,379	,005	,059	,011	,000	,079
	N	144	136	111	143	120	76	135	135

Lineare Regressionen

Körpergröße wurde als abhängige Variable über diverse Variablenkombinationen für alle Befragten, Männer und Frauen regressiert. Während der Fall 751 ($v_2=751$) der einzige univariate Ausreißer in den Modellen darstellte und jedesmal ausgeschlossen wurde, wurden daneben über die Mahalanobisdistanzen identifizierte multivariate Ausreißer in die Regressionen sowohl aus- als auch eingeschlossen.

Die erste Variablenauswahl (UV) ist hier tabelliert:

UV	Mittel	SA	N
Körpergröße	171,47	9,009	1128
logEinkomm	6,9900	,77232	1128
ALTER:	50,97	17,561	1128
Männlich	,5168	,49994	1128
Alter BL	,7881	,40882	1128
Hochschulreife	,2243	,41730	1128
Erwerbslos	,4264	,49478	1128
Ledig	,2234	,41671	1128
Verheiratet zusammenlebend	,5745	,49464	1128
20 bis 100 Tausend Einwohner	,2686	,44344	1128
Bis 20 Tausend einwohner	,4495	,49766	1128

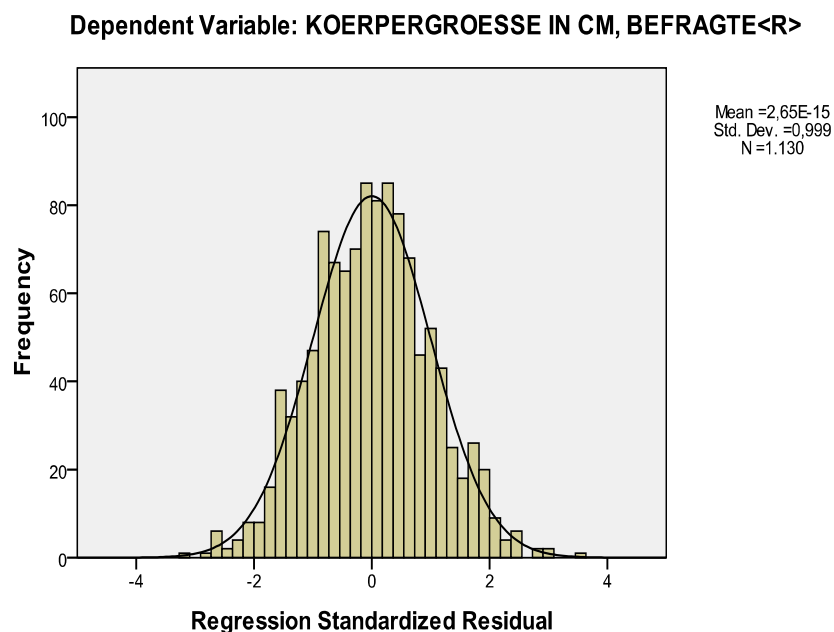
Regressiert man Körpergröße über obige Variablen und nimmt dabei die STEPWISE Methode, um das Multikollinearitätsproblem (mit allen Variablen) auszuschalten, werden folgende Modelle ausgegeben:

	Variablen	Beta	SF	St.Beta	t	Sig.	R ²	Adj. R ²
1	Konstante	165,031	,279		591,316	,000	,477	,477
	Männlich	12,451	,388	,691	32,072	,000		
2	Konstante	171,195	,606		282,428	,000	,531	,530
	Männlich	12,222	,369	,678	33,159	,000		
	Alter	-,119	,010	-,231	-11,303	,000		
3	Konstante	163,283	1,740		93,819	,000	,540	,539
	Männlich	11,622	,385	,645	30,153	,000		
	Alter	-,123	,010	-,240	-11,803	,000		
	logEinkomm	1,209	,250	,104	4,843	,000		

Hinzu wird in der ENTER-Schätzung noch der Dummy für 20 bis 100000 Einwohner signifikant ausgewiesen. Die Herausnahme von multivariaten Ausreißern (8 Fälle über extreme Mahalanobisdistanzen identifiziert) ändert daran nichts.

Die Prozedur GENLIN (plus robuste Standardfehler) wiederholt das ganze; die Residuen (für obiges 3. Modell) sind in der Grafik sichtbar:

Histogramm



Daneben wurden folgende Variablen in STEPWISE-Regressionen eingefügt:

Alter, Gewicht, Dummies für männlich, alte Bundesländer, für Hochschulreife des Vaters, für die Hochschulreife des Befragten, für Erwerbslosigkeit, Ledige, Verheiratete (zusammenlebende), Magnitudeprestige I88, Einordnung nach Terwey, Attraktivität zum Interviewsstart, Fernsehgesamtdauer, Vater: Siops I88, Alter des 1. Kindes ausser Haus, logarithmiertes Einkommen, und zuletzt Dummies für bis 20 Tausend Einwohner und zwischen 20 und 100 Tausend Einwohner.

Unter Verwendung neuer erklärenden Variablen und Multikollinearität, wird folgendes Modell (N=428, höchstes adj. $R^2 = 0,630$, ENTER-Methode) geschätzt:

Variablen	Beta	SF	St. Beta	t	Sig.
Konstante	143,815	3,445		41,748	,000
logEinkomm	1,083	,445	,090	2,435	,015
Gewicht	,204	,020	,345	10,162	,000
Männlich	8,149	,657	,468	12,402	,000
Vater Hochschul- reife	2,370	,909	,079	2,607	,009
Ledig	-1,230	1,853	-,020	-,664	,507
Alter 1. Kind ausser Haus	-,096	,023	-,128	-4,152	,000
Magnitude- prestige l88	,025	,009	,093	2,786	,006

Die Verwendung von robusten Standardfehlern zeigt ein sehr ähnliches Bild.

Männer:

Wiederholt man obengenannte Prozeduren (ENTER und STEPWISE), wird ein Modell mit den signifikanten UV: Konstante, Alter, logarithmiertes Einkommen und Hochschulreife ohne Multikollinearität (N=583, adjustiertes $R^2 = 0,146$) mit und ohne robuste Standardfehler geschätzt. Die Herausnahme von multivariaten Ausreißern erhöht das R^2 geringfügig, und lässt den Dummy für Hochschulreife in der STEPWISE-Prozedur wegfallen.

Nimmt man die zweite erweiterte Variablenkombination und Multikollinearität in Kauf, kann ein Modell mit einer Konstanten, Gewicht, logarithmiertes Einkommen, Alter des 1. Kindes ausser Haus, Dummy für Ledige und Hochschulreife des Vaters und einem adj. R^2 von 0,326 aufgestellt werden (N=223). Allein der Dummy für Ledige ist nicht signifikant (Modelle über REGRESSION / ENTER und robuste GENLIN geschätzt). Wenn man ohne 3 multivariate Fälle rechnet, wird ein etwas besser angepasstes Modell mit einer Konstanten, Gewicht, logarithmiertes Einkommen, Alter des 1. Kindes und dem Dummy für die eigene Hochschulreife vorgeschlagen.

Frauen:

Die selbe Vorgehensweise wurde für die Auswahl Frauen durchgegangen.

Ein Modell ohne Multikollinearität schätzt Körpergröße über eine Konstante, logarithmiertes Einkommen und Alter (N=546, adj. $R^2 = 0,098$). Dies bleibt so im Modell ohne multivariate Ausreißer.

Ein anderes Modell ohne MK wird mit einer Konstanten, Gewicht, Alter und einem Dummy für die Hochschulreife des Vaters gerechnet (ENTER-Methode, N=618, adj. $R^2=0,17$; Schätzungen mit und ohne robuste Fehler). Die Rechnung ohne multivariate Ausreißer bestätigt dies.

Im Vergleich zum Gesamtmodell sind die selben Modelle für Subgruppen hinsichtlich diverser Modellannahmen in den meisten Fällen schlechter gestellt (Verteilung des Störterms, Linearitätsbedingung etc.). Eine weitere Unterteilung der Gruppen Männer und Frauen wurde, wie weiter oben bei den Korrelationsmessungen, vorgenommen, jedoch nicht weiter ausgeführt. Die R^2 dieser Schätzungen liegen sowohl über (z.B. ledige Frauen; erwerbslose Frauen) als auch unter (z.B. verheiratete Frauen) jenen der zusammengefassten Modelle; daneben verlieren diverse erklärende Variablen ihre (weiter oben berichtete) Signifikanz.

Binäre logistische Regressionen

Modelle (mit Multikollinearität) mit den erklärenden Variablen Körpergröße, Alter, logarithmiertes Einkommen, Dummy für Geschlecht (männlich), Hochschulreife, Erwerbslosigkeit, Ledige und (ganztags) Beschäftigte wurden nach Ausreißereliminierung über z-Scores und ggf. über Residuen des Modells (praktisch keine) für Männer, Frauen, beide Geschlechter (jeweils plus/minus Dummy für Geschlecht) gerechnet.

Dichotome abhängige Variablen waren: alte / neue BL (v3), Internet für Info über Politik (v19), Parteipräferenz vorhanden (v69), Bildung: jeder seiner Begabung (v99), Geburt im heutigen Deutschland (v156), arbeitslos / erwerbstätig (ganztags), im öffentlichen Dienst tätig (v206), verheiratet / ledig, Mehrpersonenhaushalt (v390), Kinder/keine Kinder, leisten Sie informelle soziale Hilfe (v519), Mitglied in einer Gewerkschaft (v528), Mitglied in einer politischen Partei (v533), Arbeiter / Angestellter. Im jeweiligen Modell wurden natürlich die Dummies herausgenommen, die mit der abhängigen Variablen korrespondierten (z.B. Dummy für Ledige, AV: Verheiratet / Ledig)

Die mindestens auf 0,05 Niveau signifikante Variable Körpergröße ist tabelliert.

AV	Männer	Frauen	Alle (o. GES)	Alle (mit GES)
Ost/West				
Internetnutzung			s	
Parteipräferenz		s	s	s
Bildung				
In D v.h. geboren	s	s	s	s
Erwerbslos/-tätig				
Im öff. Dienst	s		s	
Verh. / ledig				
Mehrpers.hh.				
Eigene Kinder	s	s	s	s
Soz.Hilfe leisten?	s			
Gewerkschaftmitg.				
Parteimitglied				
Arbeiter/Angest.			s	

Ordinale logistische Regressionen

Für die abhängigen Variablen:

Hfk. von Fernsehen pro Woche, Konsumhfk. Nachrichten öff. TV, Hfk. Tageszeitung Lesen pro Woche, politisches Interesse, Inglehartindex, subjektive Schichteinstufung, allgemeine Schulbildung (in 4 Kategorien), wie viele Arbeitslose Verwandte hat der Befragte, wie häufig nutzen Sie das Internet, lieber allein oder mit anderen zusammen, Hfk. über Nacht fort, Hfk. von der Arbeit freigenommen, Vertrauen oder Vorsicht bei Kontakten, wie glücklich sind Sie, oben-unten-Skala, Zustand des Wohngebäudes, Einschätzung der Wohnumgebung, Größenklasse der politischen Gemeinde;

alle ordinal skaliert, wären ordinale logistische Regressionen geeignet. Modellschätzungen für die Gruppe der Männer, Frauen, aller Befragten plus/minus Geschlechtsdummy wurden über die unabhängigen Variablen Alter, logarithmiertes Einkommen, Dummy für Geschlecht (männlich), Hochschulreife, Erwerbslosigkeit, Ledige und (ganztags) Beschäftigte angegangen.

Die oftmals zu geringe Zellenbesetzung der nicht rekodierten abhängigen Variablen, inklusive eines ablehnenden Parallelitätstest für die übereinstimmenden Steigungskoeffizienten (Lageparameter) über

die Antwortkategorien empfahl die Wahl eines anderen Links (außer Logit), die Zusammenfassung von Kategorien oder die multinomiale logistische Regression in diesem Kontext (vgl. Janssen & Laatz, 2007: S.476).

Letztere wird ausführlicher berichtet. Die Modelle wurden alle mit Multikollinearität geschätzt. Signifikante LR-Tests der Variable Körpergröße sind tabelliert, insofern die kleinste Zelhäufigkeit zugleich auch mindestens 25 betrug. Subjektive Schichteinstufung (V120) ging hier rekodiert in die Analyse ein, ohne die Kategorien 5 und 6 (Oberschicht und keine der Schichten.)

AV	Männer	Frauen	Beide, o.G	Beide, m.G.
Zeitungslesen				s
Pol. Interesse			s	s
Schichteinstufung				s
Über Nacht fort	s			
Von der Arbeit frei			s	
Zustand Gebäude	s		s	s
Gemeindegröße	s			s

Clusteranalysen

Faktorenanalysen gingen den Clusteranalysen voraus, um Variablenreduktion über Hauptkomponentenanalyse plus Varimax-Rotation (Maximierung der erklärten Varianz, Orthogonalisierung) zu erreichen. Alle Items (Ladungen über 0,3) wurden belassen, nachdem eine vorangehende Ausschließung von Items mit Ladungen kleiner 0,5 in der nächsten HKA weitere Ausschlüsse forderte, und inhaltlich für wichtig bestimmte Variablen ausgesondert hätte. Das KMO-Kriterium (u.a.) wurde ebenfalls nicht benutzt, um schlechte Faktorenanalysen zu kennzeichnen; vielmehr wurden diverse Lebensstiloperationalisierungen benutzt, um aus diesen Gründen zu flexibilisieren.

Die hier genannten Hauptkomponentenanalysen (Blöcke 3 bis 9) entsprechen jenen im zweiten Bericht zur Studie aufgezählten (Mumdzhev, 2011). Die Tabellierung von deskriptiven Statistiken, Komponentenmatrizen etc. wird der Kürze halber erspart.

Dritter Block

Ein Beispielitem wäre hier: sollte die Frau lieber dem Mann bei der Karriere helfen.

Folgende 4 Ausprägungen sind möglich: stimme voll zu, stimme eher zu, stimme eher nicht zu, stimme gar nicht zu. Normalverteilung kann unter den Variablen anhand von Diagrammen kaum ausgemacht werden.

Das KMO beträgt 0,785; knapp 61% der Gesamtvarianz werden erklärt; N=2361.

Vierter Block

Die Items (z.B. Hfk. Politikgespräch mit Familie) haben jeweils 5 Abstufungen: sehr oft, oft, manchmal, selten, nie. Rund die Hälfte der Variablen ist annähernd normalverteilt (N=1208, KMO=0,749; 58% der Varianz werden erklärt).

Fünfter Block

Fragen wie: wer bereitet die Mahlzeiten zu im gemeinsamen Haushalt werden in diesem Block faktorisiert. Die Ausprägungen lauten: stets die Frau, meistens die Frau, hälftig gemeinsam, meistens der Mann, stets der Mann, andere Person. Schiefe Verteilungen überwiegen (N=1938, KMO=0,677; 52% der Varianz werden erklärt).

Sechster Block

Die einzelnen Variablen (z.B. Musik hören) sind mit 5 Ausprägungen, nämlich: täglich, mehrmals pro Woche, mehrmals im Monat, mehrmals im Jahr, nie, erfasst. Nur einige der Variablen sind normalverteilt (N=1349, KMO=0,756; 50% der Varianz werden erklärt).

Siebenter Block

5 Ausprägungen sind möglich: trifft voll zu, trifft eher zu, weder noch, trifft eher nicht zu, trifft gar nicht zu (z.B. schenke leicht Vertrauen, bin gutgläubig) Es sind kaum normalverteilte Variablen zu finden (N=2660, KMO=0,605; 58% der Varianz werden erklärt).

Achter Block

Die Skalen lassen folgende Antwortmöglichkeiten zu: keinen Spaß, nicht viel Spaß, etwas Spaß, großen Spaß, sehr großen Spaß (Beispiel: Spaß an Bücherlesen).

Die Variablen zumeist rechtssteil (N=1151, KMO=0,6; 66% der Varianz werden erklärt).

Neunter Block

5 Ausprägungen sind zugelassen: viel mehr, ein bisschen mehr, genauso viel, ein bisschen weniger, viel weniger (Zeit verbringen mit bezahlter Arbeit). Einige Items könnte man als normalverteilt ansehen (N=1024; KMO=0,474; knappt 65 % der Varianz werden erklärt).

Als Dimensionen von Lebensstilen wurden auch in dieser Arbeit angesehen (vgl. Klocke, 1993: S.147 ff.):

1. Freizeitverhalten und Konsummuster, welche auf das **expressive Verhalten verweisen**
2. Verkehrskreise, Zugehörigkeiten und Interaktionsmuster deuten das **interaktive Verhalten an**
3. Wertorientierungen, Einstellungen, kulturelle Traditionen und Bindungen lassen Rückschlüsse auf das **evaluative Verhalten zu**
4. Selbstwahrnehmung und Wahrnehmung der sozialen Welt ermöglichen Identitätsausbildung und zeigen das **kognitive Verhalten an**.

Die in die Lebensstiloperationalisierung eingespeisten Faktorwerte und Variablen sollten also Freizeitverhalten, Wertorientierungen und Einstellungen, das kognitive und interaktive Verhalten beschreiben. Vier Varianten bzw. Kombinationen (mit Fallzahlen zwischen 500 und 700) wurden in Clusteranalysen vorgenommen:

1. Faktorwerte der Blöcke 4,6,8,9
2. Faktorwerte der Blöcke 3,5,6,8
3. Faktorwerte der Blöcke 3,5,7,8
4. Faktorwerte der Blöcke 3,4,7,9

Desweiteren wurden folgende Variablen (als intervallskaliert angenommen) in ihrer ursprünglichen Kodierung für jede der oben genannten Faktorwertekombination verwendet:

Variable	Ausprägungen
HÄUFIGKEIT VON FERNSEHEN PRO WOCHE	NIE SELTENER AN EINEM TAG AN 2 TAGEN AN 3 TAGEN AN 4 TAGEN AN 5 TAGEN AN 6 TAGEN AN ALLEN 7 TAGEN
FERNSEHGESAMTDAUER PRO TAG IN MINUTEN	stetig
KONSUMHÄUFIGK.: NACHRICHTEN	NIE

OEFF. TV	SELTENER AN EINEM TAG AN 2 TAGEN AN 3 TAGEN AN 4 TAGEN AN 5 TAGEN AN 6 TAGEN AN ALLEN 7 TAGEN
KONSUMHÄUFIGK.: NACHRICHTEN PRIVATES TV	NIE SELTENER AN EINEM TAG AN 2 TAGEN AN 3 TAGEN AN 4 TAGEN AN 5 TAGEN AN 6 TAGEN AN ALLEN 7 TAGEN
HÄUFIGKEIT TAGESZEITUNG LESEN PRO WOCHE	NIE SELTENER AN EINEM TAG AN 2 TAGEN AN 3 TAGEN AN 4 TAGEN AN 5 TAGEN AN 6 TAGEN AN ALLEN 7 TAGEN
BEWERTUNG VON DEMOKRATIE ALS IDEE	SEHR DAFUER ZIEMLICH DAFUER ETWAS DAFUER ETWAS DAGEGEN ZIEMLICH DAGEGEN SEHR DAGEGEN
POLITISCHES INTERESSE, BEFR. <ORDINAL>	SEHR STARK STARK MITTEL WENIG UEBERHAUPT NICHT
WICHTIGKEIT VON RUHE UND ORDNUNG	AM WICHTIGSTEN AM ZWEITWICHTIGSTEN AM DRITTWICHTIGSTEN AM VIERTWICHTIGSTEN
WICHTIGKEIT VON BÜRGEREINFLUSS	AM WICHTIGSTEN AM ZWEITWICHTIGSTEN AM DRITTWICHTIGSTEN AM VIERTWICHTIGSTEN
WICHTIGKEIT DER INFLATIONSBEKÄMPFUNG	AM WICHTIGSTEN AM ZWEITWICHTIGSTEN AM DRITTWICHTIGSTEN AM VIERTWICHTIGSTEN
WICHTIGKEIT V. FREIER MEINUNGSÄUSSERUNG	AM WICHTIGSTEN AM ZWEITWICHTIGSTEN AM DRITTWICHTIGSTEN AM VIERTWICHTIGSTEN
LINKS-RECHTS-SELBSTEINSTUFUNG	1 LINKS-...-10 RECHTS
GERECHTER ANTEIL A.LEBENSSTANDARD,BEFR.?	SEHR VIEL WENIGER ETWAS WENIGER GERECHTEN ANTEIL MEHR ALS GERECHTEN

NACHBARSCHAFTSKONTAKT: HÄUFIGKEIT	KEIN KONTAKT SELTEN KONTAKT GELEGENTLICH KONTAKT HAEUFIG KONTAKT SEHR HAEUFIG KONTAKT
ZEITAUFWAND FÜR VEREINSKOLLEGEN	TAEGLICH 1X JEDE WOCHE 1X JEDEN MONAT SELTENER NIE
ZEITAUFWAND FÜR ARBEITSKOLLEGEN	TAEGLICH 1X JEDE WOCHE 1X JEDEN MONAT SELTENER NIE
ZEITAUFWAND FÜR FREUNDE	TAEGLICH 1X JEDE WOCHE 1X JEDEN MONAT SELTENER NIE
VERTRAUEN ODER VORSICHT BEI KONTAKTEN	FAST IMMER VERTRAUEN NORMALERW. VERTRAUEN NORMALERW. VORSICHT FAST IMMER VORSICHT

Alle Variablen wurden standardisiert (als z-Scores) in die Clusteranalyse eingesetzt.

Das single-linkage Verfahren (über quadrierte euklidische Distanz) wurde für 5-10 Clusterlösungen gerechnet, und die resultierenden 1-Fall Cluster als Ausreißer interpretiert, und von nachfolgenden Ward-Clusteranalysen ausgeschlossen (für einige der folgenden multinomialen logistischen Regressionen wurden dadurch bestimmte Mindestgrößen der Cluster bzw. Zellhäufigkeiten unterschritten). Parallel dazu lieferte die kmeans-Analyse zu unterschiedliche Clustergrößen, die für die weiteren Rechnungen nicht zu verwenden waren. Auf inhaltliche Beschreibungen der Cluster, sowie auf Argumentationen über die optimale Zahl der Cluster wurde hier verzichtet.

Die WARD-Clusterlösungen wurden als Faktoren einer Varianzanalyse der Variable Körpergröße eingespeist, desweiteren wurden sie in multinomialen logistischen Regressionen über Variablen der Sozialstruktur inklusive Körpergröße regressiert. Erklärende Variablen waren explizit: Dummy für Geschlecht (männlich), Dummy für Hochschulreife, Dummy für Ledige, Dummy für Verheiratete, Dummy für Erwerbslose, Alter, logarithmiertes Einkommen, sowie die Variable Körpergröße.

Die Signifikanzen der Streuungserlegung (F-Tests bzw. robuste Alternativen) und die LR-Tests der Regressionen sind tabelliert.

Diskriminanzanalysen haben die Ergebnisse der logistischen Regressionen (mit Multikollinearität, bei höheren Clusterlösungen mit Fallzahlen unter 25 bzw. mit rechnerischen Problemen) bestätigt, und maßen der Variable Körpergröße ebenfalls geringe (absolute und relative) Bedeutung zu.

Erste Lebensstiloperationalisierung

	5 Cluster	6 Cluster	7 Cluster	8 Cluster	9 Cluster	10 Cluster
VA männlich	s	s	s	s	s	s
VA Weiblich						
VA alle	s		s			
MLR, mit Geschlecht		s				s
MLR, ohne Geschlecht						

Zweite Lebensstiloperationalisierung

	5 Cluster	6 Cluster	7 Cluster	8 Cluster	9 Cluster	10 Cluster
VA männlich						
VA Weiblich						
VA alle	s	s		s		
MLR, mit Geschlecht				s	s	s
MLR, ohne Geschlecht						

Dritte Lebensstiloperationalisierung

	5 Cluster	6 Cluster	7 Cluster	8 Cluster	9 Cluster	10 Cluster
VA männlich						
VA Weiblich	s	s			s	s
VA alle	s	s	s	s	s	s
MLR, mit Geschlecht	s				s	
MLR, ohne Geschlecht						

Vierte Lebensstiloperationalisierung

	5 Cluster	6 Cluster	7 Cluster	8 Cluster	9 Cluster	10 Cluster
VA männlich			s			
VA Weiblich						
VA alle		s	s	s	s	s
MLR, mit Geschlecht	s	s	s	s	s	s
MLR, ohne Geschlecht			s	s	s	s

Verwendete Literatur

Janssen, Jorgen & Laatz, Wilfried 2007: Statistische Datenanalyse mit SPSS für Windows, Berlin: Springer.

Klocke, Andreas 1993: Sozialer Wandel, Sozialstruktur und Lebensstile in der Bundesrepublik, Deutschland, Frankfurt: Peter Lang.

Mumdzhev, Milko 2011: Zweiter Zwischenbericht zur Studie: Körpergröße als Indikator sozialer Ungleichheit, Nürnberg: Arbeitspapier.